**Supplementary Materials**

***Study design and study population***

## The current study consists of an initial cross-sectional screening phase followed by a prospective, longitudinal, parallel-group phase. The study was conducted in several Egyptian centers in Cairo and Upper Egypt (Ain Shams University Hospitals, Cairo, and Gastroenterology centers in Cairo and Minya) from September, 2011 through July, 2016. The study protocol and patients’ informed consent were approved by the institutional review boards and independent ethics committees at the participating sites. The study was conducted in accordance with the Declaration of Helsinki and was consistent with the International Conference on Harmonization and Good Clinical Practice.

## Patients with hepatic steatosis (mild, moderate or severe 33,34)detected by ultrasound (Philips EPIQ7G ultrasound machine; Philips, Reedsville, PA, USA) suggestive of NAFLD/NASH were further tested by “NAFLD liver fat score (NAFLD-LFS)”,35 “Fatty Liver Index (FLI)” 36 and “Hepatic Steatosis Index (HSI)” 37 were then calculated according to the previously described formulas. *(Supplementary material)*

NAFLD-LFS=2.89 +1.18xMS (yes=1/no=0) +0.45xT2DM (yes=2/no=0)+0.15xI0+ 0.04 xAST-0.94 x AST/ALT.I0 (µU/ml) represents fasting insulin and AST represents fasting AST levels (U/l). Values ≤−0.640 rule out, while values >−0.640 rule in NAFLD.35

FLI = Logistic (0.953x*In* (TG) +0.139 x BMI+o.718 + In (gGT) x waist -15.745 x100. Values <30 rule out and values ≥60 rules in steatosis. 36  logistic(x) = 1/(1+e-x) denotes the logistic function and *ln* the natural logarithm.36

HSI: 8xALT/AST ratio +BMI+2 (if diabetic) +2 (if female). Values <30 rule out steatosis and values >36 ruling in steatosis.37

## Patients were subjected to specific investigations to exclude alcoholic liver disease, viral hepatitis (HBV and HCV), Wilson disease, haemochromatosis, autoimmune hepatitis and drug related steatosis. Patients were enrolled in the study if they fulfilled the following criteria: i) presence of hepatic steatosis; ii) NAFLD-LFS values >−0.640, FLI values > 60 and HSI values >36 iii) absence of any evidence of other chronic liver diseases and other causes of hepatic steatosis, iii) no history of significant alcohol consumption, iv) elevated aminotransferase levels found in one of three situations. Patients fulfilling the inclusion criteria were further evaluated for the stage of liver disease by transient elastography (TE). Serial TE examinations (Fibroscan® , Echosens, Paris, France) were performed at baseline and during follow up as previously described.38  The liver stiffness results were reported in kilopascals (kPa). The LSM measurement was classified into F0/1, F1 ± 1, F1/2, F2/3, F3 ± 1, F3/4, F4. The LSM values were used to estimate the METAVIR fibrosis stage as follows: F0/F1: 2-4.5; F1±1; 4.6 to 6 kPa; F1/2: 6.1 to 8.7 kPa; F2/3: 8.8 to 11kPa; F3±1: 12-18 kPa; F3/4: 18-38.6 kPa and F4 >38.6-75 kPa. 38

**Screening for celiac disease**

Patients with NAFLD/NASH provided written informed consents before the screening phase of the study which included completing validated questionnaires and screening by tissue transglutaminase antibodies (TTGA; QUANTA Lite human-TTGA ELISA kit ; INOVA Diagnostics, San Diego, CA, USA) according to the manufacturer recommendations. Patients with symptoms suggestive of celiac disease but negative TTGA were tested for potential IgA deficiency (Abcam human IgA ELISA Kit). Patients with IgA deficiency were screened by deamidated gliadin peptide IgG antibody (DGP) *(*DGP IgG ELISA kit; Biosource, San Diego, CA,    USA). Patients with positive TTGA or DGP were informed that they may have celiac disease and were invited to join the study and undergo further investigations. Those who accepted signed another informed consent before entry and before any study related investigation or upper endoscopy.

1. **Questionnaires:**
2. ***Coeliac UK assessment tool***: At enrollment, patients with hepatic steatosis were invited to complete an Arabic version of the questionnaire adapted from the Coeliac UK assessment tool available at: <https://www.isitcoeliacdisease.org.uk/login?redirect=assessment>. 40 The responses were analyzed and recommendations for further CeD testing were made according to the guidelines of the National Institute of Health and Care Excellence (NICE) available at <https://www.nice.org.uk/guidance/ng20>.
3. **Gastrointestinal Symptom Rating Scale** (CeDGSRS) questionnaire (Supplementary material).:47 Patients with positive TTGA and EMA and patients with high suspicion of CeD who were enrolled in the second phase of the study completed an Arabic version of Gastrointestinal Symptom Rating Scale (GSRS), An English version is shown below
4. A locally validated food-frequency Arabic questionnaire (FFQ) was used to capture food intake and dietary intake over 7 days. The questionnaire contained several local food items in seven broad categories: bread / pasta / rice; vegetables; fruit; meat / fish / egg; ol, fat, butter, beverages; snacks; soups; and salt / sauces. Subjects were asked to complete the questionnaires under the supervision of a trained research staff, with food models, food containers, and a catalogue of pictures of individual food portions provided to facilitate portion size estimation. The amount of cooking oil was calculated based on the usual cooking methods, the usual type of cooking oil and the usual portion of different foods used by the subjects, Daily nutrient intakes and food group intakes were estimated using Food Processor Nutrition analysis and Fitness software version 8.0 (ESHA Research, Salem, Oregon, USA).
5. **Gluten-free diet compliance questionnaire (GFDCQ):** 51An Arabic version ofthis questionnaire which monitors adherence to gluten free diet is shown below.

Appendix A: **Gastrointestinal Symptom Rating Scale** (CeDGSRS) questionnaire

Name:

A rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Circle the number which best represents the current severity of the symptom.

**1. Abdominal pains.** Representing subjectively experienced bodily discomfort, aches and pains.

The type of pain may be classified according to the patient's description of the appearance and quality of the pain as epigastric, on the basis of typical location, association with acid-related symptoms, and relief of pain by food or antacids; as colicky when occurring in bouts, usually

with a high intensity, and located in the lower abdomen; and as dull when continuous, often for several hours, with moderate intensity.

Rate according to intensity, frequency, duration, request for relief, and impact on social performance.

0 No or transient pain

1 Occasional aches and pains interfering with some social activities

2 Prolonged and troublesome aches and pains causing requests for relief and interfering with many social activities

3 Severe or crippling pains with impact on all social activities

**2. Heartburn.** Representing retrosternal discomfort or burning sensations. Rate according to intensity, frequency, duration, and request for relief.

0 No or transient heartburn

1 Occasional discomfort of short duration

2 Frequent episodes of prolonged discomfort; requests for relief

3 Continuous discomfort with only transient relief by antacids

**3. Acid regurgitation.** Representing sudden regurgitation of acid gastric content. Rate according to intensity, frequency, and request for relief.

0 No or transient regurgitation

1 Occasional troublesome regurgitation

2 Regurgitation once or twice a day; requests for relief

3 Regurgitation several times a day; only transient and insignificant relief by antacid**4. Sucking sensations in the epigastrium.** Representing a sucking sensation in the epigastrium with relief by food or antacids. If food or antacids are not available, the sucking sensations progress to ache, and pains. Rate according to intensity, frequency, duration, and request for relief.

0 No or transient sucking sensation

1 Occasional discomfort of short duration; no requests for food or antacids between meals

2 Frequent episodes of prolonged discomfort, requests for food and antacids between meals

3 Continuous discomfort; frequent requests for food or antacids between meals

**5. Nausea and vomiting.** Representing nausea which may increase to vomiting. Rate according to intensity, frequency, and duration.

0 No nausea

1 Occasional episodes of short duration

2 Frequent and prolonged nausea; no vomiting

3 Continuous nausea; frequent vomiting

**6. Borborygmus.** Representing reports of abdominal rumbling. Rate according to intensity, frequency, duration, and impact on social performance

0 No or transient borborygmus

1 Occasional troublesome borborygmus of short duration

2 Frequent and prolonged episodes which can be mastered by moving without impairing social performance

3 Continuous borborygmus severely interfering with social performance

**7. Abdominal distension.** Representing bloating with abdominal gas. Rate according to intensity, frequency, duration, and impact on social performance.

0 No or transient distension

1 Occasional discomfort of short duration

2 Frequent and prolonged episodes which can be mastered by adjusting the clothing

3 Continuous discomfort seriously interfering with social performance

**8. Eructation.** Representing reports of belching. Rate according to intensity, frequency, and impact on social performance.

0 No or transient eructation

1 Occasional troublesome eructation

2 Frequent episodes interfering with some social activities

3 Frequent episodes seriously interfering with social performance

**9. Increased flatus.** Representing reports of excessive wind. Rate according to intensity, frequency, duration, and impact on social performance

0 No increased flatus

1 Occasional discomfort of short duration

2 Frequent and prolonged episodes interfering with some social activities

3 Frequent episodes seriously interfering with social performance

**10. Decreased passage of stools.** Representing reported reduced defecation. Rate according to frequency. Distinguish from consistency.

0 Once a day

1 Every third day

2 Every fifth day

3 Every seventh day or less frequently

**11. Increased passage of stools.** Representing reported increased defecation. Rate according to frequency. Distinguish from consistency.

0 Once a day

1 Three times a day

2 Five times a day

3 Seven times a day or more frequently

**12. Loose stools.** Representing reported loose stools. Rate according to consistency independent of frequency and feelings of incomplete evacuation.

0 Normal consistency

1 Somewhat loose

2 Runny

3 Watery

**13. Hard Stools.** Representing reported hard stools. Rate according to consistency independent of frequency and feelings of incomplete evacuation.

0 Normal consistency

1 Somewhat hard

2 Hard

3 Hard and fragmented, sometimes in combination with diarrhea

**14. Urgent need for defecation.** Representing reports of urgent need for defecation, feelings of incomplete control, and inability to control defecation. Rate according to intensity, frequency, and impact on social performance.

0 Normal control

1 Occasional feelings of urgent need for defecation

2 Frequent feelings of urgent need for defecation with sudden need for a toilet interfering with social performance

3 Inability to control defecation

**15. Feeling of incomplete evacuation.** Representing reports of defecation with straining and a feeling of incomplete evacuation of stools. Rate according to intensity and frequency.

0 Feeling of complete evacuation without straining

1 Defecation somewhat difficult; occasional feelings of incomplete evacuation

2 Defecation definitely difficult; often feelings of incomplete evacuation

3 Defecation extremely difficult; regular feelings of incomplete

**Patients assessment**

Patients with positive serology were subjected to careful history, clinical examination, laboratory investigations including urine and stools analysis, complete blood picture, lipid profile study, liver functions, fasting insulin and fasting glucose, homeostasis model assessment-insulin resistance (HOMA-IR),41 serum iron, ferritin, folic acid, vitamins D and B12, antinuclear antibodies, thyroid function tests, cytokine assessment and gastrointestinal endoscopy with duodenal biopsy. HLA-DQ typing for the presence of DQB1\*02 and DQB1\*0302 ( (PCR sequence-specific oligonucleotide typing (QIAxcel system and QIAxcel DNA Fast Analysis Kit Product # 929008, QIAGEN, Stamford, CT, USA) was performed in a subset of patients.

**Cytokines assessment**

Tumor necrosis factor-alpha (TNF- ELISA kit;BioSource, San Diego, CA,    USA), **interleukin 1 (IL-1 ELISA Kit;** BioSource, San Diego, CA,    USA), interleukin 6 (human IL- 6 ELISA Kit, , BioSource, San Diego, CA,    USA), Interleukin 15 (IL-15 ELISA Kit;BioSource, San Diego, CA,    USA), interleukin 17 (IL-17 ELISA Kit ;BioSource, San Diego, CA, USA;), human interleukin 10 (IL-10 ELISA kit, BioSource, San Diego, CA, USA), tumor growth factor beta 1 (human TGF beta 1 ELISA kit, BioSource, San Diego, CA,  USA) and YKL-40 (human YKL-40 ELISA kit, Quidel, San Diego, CA, USA) were measured at baseline and end of follow-up according to the manufacturers’ instructions.

**Endoscopy and small intestinal biopsy**

At baseline, all patients with serologic evidence of CeD had upper endoscopy and duodenal biopsy which were repeated one year after GFD in a subset of patients (101 patients in group A and 30 patients in group C). At least 6 mucosal biopsies were taken from the second part of duodenum and bulb. Sections were stained with hematoxylin and eosin-stained and Giemsa and were examined by an experienced gastrointestinal pathologist (author: L.N.) according to the Modified Marsh classification for celiac disease 42 which assesses the intraepithelial lymphocytes per 100 enterocytes (IEL/100 enterocytes), crypt hyperplasis and villi. IELs reference value was set at <40 cells/100 epithelial cells. IEL/100 enterocytes > 40, increased crypt hyperplasia, marked or complete villous atrophy and mucosal villous height-crypt depth ratio (VH/CrD) <2.0 were indicative of active celiac disease. 42

According to clinical manifestations, serology and intestinal biopsy, patients with celiac were classified into: i) Symptomatic CeD with manifestations related to celiac with positive serology and histologic manifestations in intestinal biopsy; ii) Silent CeD with no or minimal symptoms, “damaged” mucosa and positive serology and iii) Latent CeD with positive serology but with normal intestinal mucosa and no symptoms.16,17

**Initiation of GFD and assessment of efficacy**

Patients diagnosed with CeD were informed about the disease and the importance and benefits of following a lifelong GFD. A nutritional consultation and detailed information sheet including gluten free food items was discussed and provided to all patients. Adherence to GFD was assessed during clinical visits scheduled every 3 months. During each visit, patients completed a validated structured questionnaire that evaluated their adherence and symptoms and had clinical examination. After 1 year of GFD, patients repeated serological assays, follow up biopsy, abdominal ultrasound and TE. Complete clinical improvement was defined as the complete resolution of baseline symptoms after 1 year of GFD. Clinical partial improvement was defined as a resolution of at least 50% of the baseline symptoms after 1 year of GFD. Complete histological improvement is defined as resolution of villous atrophy associated with the absence of crypt hyperplasia and ≤40/100 intraepithelial lymphocytes. Partial histological recovery is defined as improvement of at least one grade on the Marsh classification compared with the initial histology.